ELSEVIER

Statistics & Probability Letters 37 (1998) 409–414

STATISTICS & PROBABILITY LETTERS

Addition or deletion?

Aloke Dey[a,*], Chand K. Midha[b]

[a] *Indian Statistical Institute, New elhi 110 016, India*

[b] *epartment of Mathematical Sciences, The University of Akron, Akron, OH 44325, USA*

Received 1 April 1997; received in revised form 1 August 1997

Abstract

Suppose it is desired to have an optimal' resolution III fraction of a 2^p factorial in N runs where $N \equiv 2 (\mathrm{mod}\, 4)$. A design for this purpose can be obtained by adding two runs optimally to the $n \times p$ matrix derived by a suitable choice of p columns of H_n, a Hadamard matrix of order n. Alternatively, one can think of deleting two runs in an optimal manner from the $(n+4) \times p$ matrix derived from H_{n+4}. A natural question then arises: do these two strategies give designs that are equally efficient in terms of a well defined optimality criterion? We show that for $p = 2$ or 3, the design obtained by deletion is as good as the addition design under the A- or the D-optimality criterion. However, for $p \geqslant 4$, the performance of the deletion design compared to the optimal addition design is rather poor as per the D-criterion, especially for large values of p. Under the A-criterion, the addition design is always better than the deletion design for $p \geqslant 4$, but the loss of efficiency using the deletion design is not too large for moderate values of p.

AMS classification: 62K15

Keywords: Resolution III fractions; Optimality

1. Introduction and preliminaries

A fractional factorial design is said to be of resolution III if it allows the estimability of the mean and all main effects under the assumption that all interactions involving two or more factors are negligible. In this paper, we consider resolution III fractions of 2^p factorials. We assume that the *Hadamard conjecture* is true, i.e., there exists a Hadamard matrix of order $n > 2$ whenever $n \quad 0 (\mathrm{mod}\, 4)$. A positive integer $n \quad 0 (\mathrm{mod}\, 4)$ will be called a Hadamard number. A Hadamard matrix of order n will be denoted by H_n and we shall assume (without loss of generality) that the first column of H_n consists of only $+1$'s.

Suppose it is desired to have a resolution III fraction for a 2^p factorial in $N \quad 2 (\mathrm{mod}\, 4)$ runs, which is optimal in some sense. A design for this purpose can be obtained by first deleting the first column of all ones from an H_n or H_{n+4} and retaining any p columns of the remaining columns to get an $n \times p$ or $(n+4) \times p$ matrix and then either (i) adding two runs optimally' to the $n \times p$ matrix derived from H_n, or, (ii) deleting

* Corresponding author.

0167-7152/98/$19.00
PII S0167-7152(97)00144-2

two runs optimally from the $(n+4)\times p$ matrix derived from H_{n+4}. Do the procedures (i) and (ii) give rise to designs that are equally efficient according to some well defined optimality criterion? In this paper, we attempt to answer this question with respect to the two commonly used optimality criteria, viz., the A- and the D-criterion.

Cheng (1980), among other things, showed that adding a single run to a 2-symbol orthogonal array of strength $2u$ with $m-1$ runs or, deleting a run from a 2-symbol orthogonal array of strength $2u$ with $m+1$ runs gives an m-run resolution-$(2u+1)$ design for a two-level factorial that is optimal according to a wide class of criteria. Mitchell (1974) while discussing his DETMAX algorithm for finding D-optimal fractions of two level factorials of resolution III suggested that a D-optimal fraction with N 2 (mod 4) may be obtained by adding two runs to an orthogonal design with $N-2$ runs. See also Payne (1974), who considers the problem of maximizing the determinant of $A'A$ where A is an $n\times p$ matrix with entries 1.

Let X_0 denote the $n\times(p+1)$ design matrix corresponding to the resolution III fraction of a 2^p factorial in n 0 (mod 4) runs. The columns of X_0 correspond to the mean and p main effects. Then, it is easy to verify that $X_0'X_0=nI_{p+1}$, where I_m denotes an identity matrix of order m. Let two more runs be added to the design in n runs, and we call the new design in $n+2$ runs an addition' design. Let the $2\times(p+1)$ matrix of the two added rows of the new design matrix be denoted by X_1, that is, the design matrix of the design with $n+2$ runs, say X_a is

$$X_\mathrm{a}=\begin{pmatrix}X_0\\X_1\end{pmatrix},$$

so that

$$X_\mathrm{a}'X_\mathrm{a}=X_0'X_0+X_1'X_1=nI_{p+1}+X_1'X_1.$$

The eigenvalues of $X_\mathrm{a}'X_\mathrm{a}$ are therefore $n+\lambda_i$, where for $i=1,2,\ldots,p+1$, λ_i are the eigenvalues of $X_1'X_1$. Since the nonzero eigenvalues of $X_1'X_1$ and those of X_1X_1' are identical, it is easier to work with the 2×2 matrix X_1X_1'. Let the added runs, each with two distinct entries, differ at t coordinates. Then, it can be seen that

$$X_1X_1'=\begin{pmatrix}p+1 & p+1-2t\\ p+1-2t & p+1\end{pmatrix}.$$

The eigenvalues of X_1X_1' are $2t$ and $2(p+1-t)$. Hence the eigenvalues of $X_\mathrm{a}'X_\mathrm{a}$ are n with multiplicity $p-1$, $n+2t$ and $n+2(p+1-t)$.

Now consider a design for a 2^p factorial in $n+4$ runs derived from H_{n+4}. We delete two runs from this design to get a design for a 2^p factorial in $n+2$ runs and call this design a deletion' design. Let the design matrix of the $(n+4)$-run design be denoted by X_2 and let X_3 denote the design matrix corresponding to the two deleted runs. If X_d denotes the design matrix of the deletion design with $n+2$ runs, then

$$X_2=\begin{pmatrix}X_\mathrm{d}\\X_3\end{pmatrix},$$

so that

$$X_\mathrm{d}'X_\mathrm{d}=X_2'X_2-X_3'X_3=(n+4)I_{p+1}-X_3'X_3.$$

Hence, the eigenvalues of $X_\mathrm{d}'X_\mathrm{d}$ are $(n+4)-\mu_i$, where for $i=1,2,\ldots,p+1,\mu_i$ are the eigenvalues of $X_3'X_3$. Arguing as before, we therefore have that the eigenvalues of $X_\mathrm{d}'X_\mathrm{d}$ are $n+4$ with multiplicity $(p-1)$, $n+4-2t$ and $n+4-2(p+1-t)$.

2. Comparison based on the A criterion

Let $\theta_1, \theta_2, \ldots, \theta_{p+1}$ be the eigenvalues of the information matrix $X_D'X_D$ of a resolution III fraction D of a 2^p factorial. Then, the A-criterion requires the minimization of $A = \theta_1^{-1} + \cdots + \theta_{p+1}^{-1}$. From our discussion in the previous section, it follows that the value of the A-criterion for the addition design, as a function of t is given by

$$A_{\mathrm{a}}(t) = \frac{p-1}{n} + \frac{1}{n+2t} + \frac{1}{n+2p+2-2t}.$$

If p is odd, the minimum of $A_{\mathrm{a}}(t)$ occurs at $t = (p+1)/2$. The minimum of $A_{\mathrm{a}}(t)$, which we denote by $A_{\mathrm{a}}(O)$, is given by

$$A_{\mathrm{a}}(O) = \frac{(p-1)}{n} + \frac{2}{(n+p+1)} \quad \text{if } p \text{ is odd.} \tag{2.1}$$

When p is even, the minimum of $A_{\mathrm{a}}(t)$ is

$$A_{\mathrm{a}}(O) = \frac{p-1}{n} + \frac{1}{n+p} + \frac{1}{n+p+2} \quad \text{if } p \text{ is even.} \tag{2.2}$$

For the deletion design, the minimum of the A-criterion, denoted by $A_{\mathrm{d}}(O)$ are given by

$$A_{\mathrm{d}}(O) = \frac{p-1}{n+4} + \frac{2}{n-p+3} \quad \text{if } p \text{ is odd;} \tag{2.3}$$

$$A_{\mathrm{d}}(O) = \frac{p-1}{n+4} + \frac{1}{n-p+4} + \frac{1}{n-p+2} \quad \text{if } p \text{ is even.} \tag{2.4}$$

If p is *odd*, we have

$$\begin{aligned} A_{\mathrm{d}}(O) - A_{\mathrm{a}}(O) &= \frac{p-1}{n+4} + \frac{2}{n-p+3} - \frac{p-1}{n} - \frac{2}{n+p+1} \\ &= \frac{4(p-1)(p^2-2p-3)}{n(n+4)(n-p+3)(n+p+1)}. \end{aligned}$$

Clearly, $A_{\mathrm{d}}(O) \geqslant A_{\mathrm{a}}(O)$ for all $p \geqslant 3$, with equality if and only if $p = 3$. We thus have

Theorem 2.1. *If $p > 3$ is odd, the best addition design is superior to the best deletion design on the basis of the A-optimality criterion. For $p = 3$, both the designs are equally efficient as per the A-criterion.*

If p is even, we have

$$\begin{aligned} &A_{\mathrm{d}}(O) - A_{\mathrm{a}}(O) \\ &= \frac{p-1}{n+4} + \frac{1}{n-p+4} + \frac{1}{n-p+2} - \frac{p-1}{n} - \frac{1}{n+p} - \frac{1}{n+p+2} \\ &= \frac{2(p-2)\ N^2(p-4) + N(p^2+6p-16) - 2p^3 + 6p^2 + 12p - 16}{n(n+4)(n-p+4)(n^2+4n+4-p^2)}. \end{aligned}$$

Clearly, $A_d(O) = A_a(O)$ if $p=2$. For $p>2$, it can be seen that $A_d(O)>A_a(O)$. Hence, we have

Theorem 2.2. *If $p > 2$ is even, the best addition design is superior to the best deletion design on the basis of the A-optimality criterion. For $p=2$, both the designs are equally efficient as per the A-criterion.*

In order to see how the best deletion design compares with the best addition design with respect to the A-criterion, the values of $e_1 = A_a(O)/A_d(O)$ was computed for all Hadamard numbers n in the interval $[4,48]$ and for all $4 \leqslant p \leqslant n-1$. It turns out that the values of e_1 range between 99.9 $(n=32, p=4,5;\ n=36,40,\ 4 \leqslant p \leqslant 7;\ n=44,48, 4 \leqslant p \leqslant 8)$ to 70 $(n=48, p=47)$. Thus, the deletion design is nearly as good as the addition design for moderate values of p. A graph showing the values of e_1 for $8 \leqslant n \leqslant 48$ and $2 \leqslant p \leqslant n-1$ is given in Fig. 1.

3. Comparison based on the D criterion

Recall that a design D is D-optimal if and only if D maximizes $\prod_{i=1}^{p+1} \theta_i$, where as before, $\theta_1, \ldots, \theta_{p+1}$ are the eigenvalues of the information matrix $X_D'X_D$ of D. Let $D_a(t)$ and $D_d(t)$, respectively, denote the value of the D-criterion for the addition and deletion designs, as a function of t. Then,

$$D_a(t) = n^{p-1}(n+2t)(n+2p+2-2t),$$

$$D_d(t) = (n+4)^{p-1}(n+4-2t)(n+4-2p-2+2t).$$

The maximum values of $D_a(t)$ and $D_a(t)$ are given by

$$\begin{aligned} D_a(O) &= n^{p-1}(n+p+1)^2, \\ D_d(O) &= (n+4)^{p-1}(n-p+3)^2 \quad \text{if } p \text{ is odd}, \end{aligned} \tag{3.1}$$

$$\begin{aligned} D_a(O) &= n^{p-1}(n+p)(n+p+2), \\ D_d(O) &= (n+4)^{p-1}(n-p+4)(n-p+2) \quad \text{if } p \text{ is even}. \end{aligned} \tag{3.2}$$

The expressions for $D_a(O)$ for both even and odd p are identical to the maximal determinant values of $A'A$ where A is an $N \times m$ matrix with entries 1, as given by Payne (1974). Therefore, the best addition design is indeed D-optimal and we have

Theorem 3.1. *The best addition design is a D-optimal resolution III fraction of a 2^p factorial in $n+2$ runs, where n is a Hadamard number.*

To see how the best deletion design fares in comparison to the D-optimal addition design, the expressions of $D_a(O)$ and $D_d(O)$ were numerically evaluated for $4 \leqslant p \leqslant n-1$ and all Hadamard numbers n in the interval $[4,48]$. It is easy to verify that for $p=2$ or 3, both the strategies are equally good. The efficiency' of the deletion design with respect to the addition design, as measured by the ratio $D_d(O)/D_a(O) = e_2$, say, decreases monotonically with p for each of the values of n. The value of e_2 is at least 90 for $4 \leqslant p < n/2$, but once p exceeds $n/2$, the values of e_2 fall sharply for moderate values of n. As n increases, the fall in the values of e_2 is however, not very rapid. A graph showing the values of e_2 for $8 \leqslant n \leqslant 48$ and $2 \leqslant p \leqslant n-1$ is given in Fig. 1.

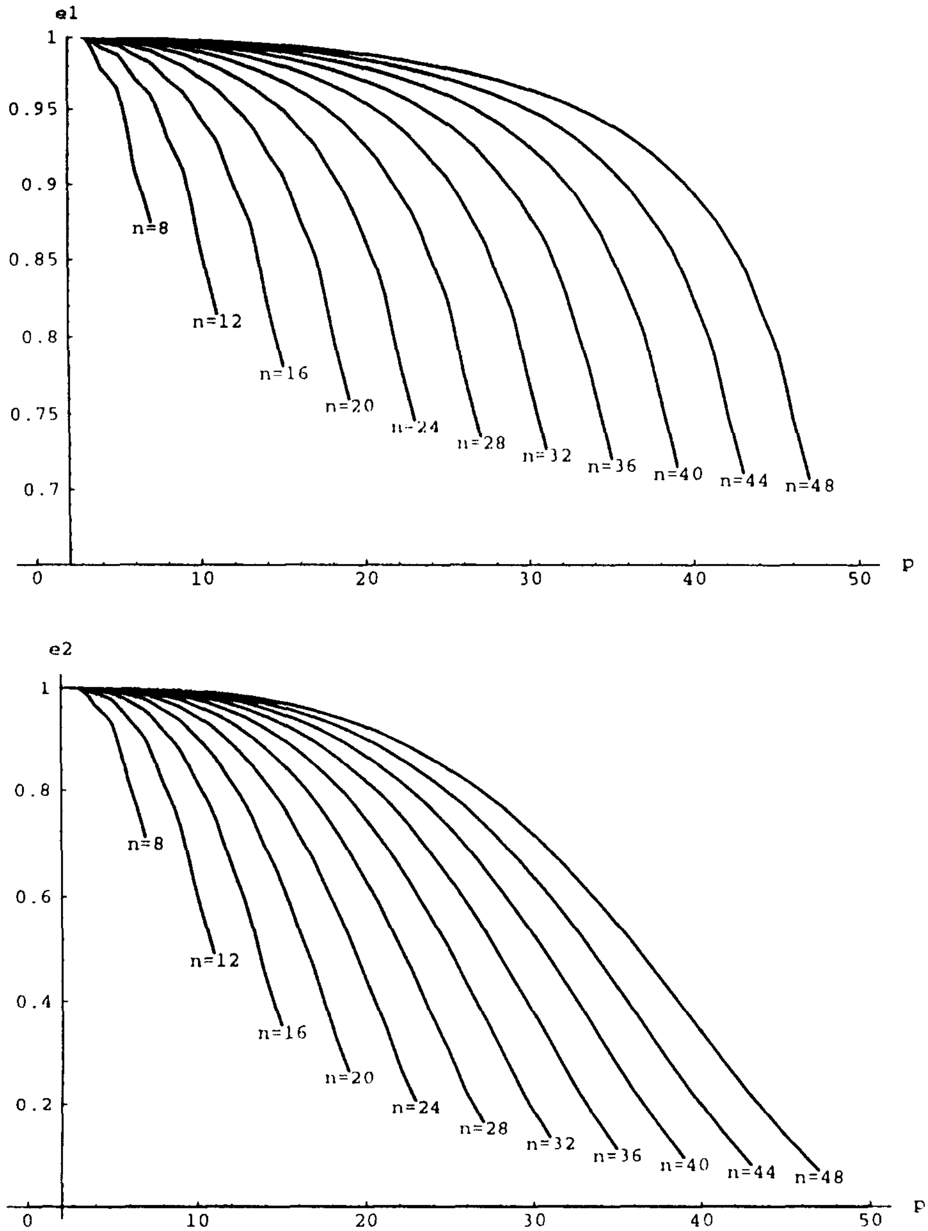

Fig. 1. e_1 and e_2-values for various values of n and p.

Acknowled ements

The authors would like to thank a referee for useful comments on a previous draft. Thanks are also due to D. Stark and Arupkumar Pal for their help with the computations and preparation of the graphs.

References

Cheng, C.-S., 1980. Optimality of some weighing and 2^n fractional factorial designs. Ann. Statist. 8, 436–446.
Mitchell, T.J., 1974. Computer construction of "D-optimal" first order designs. Technometrics 16, 211–220.
Payne, S.E., 1974. On maximizing det $(A^{\mathrm{T}}A)$. Discrete Math. 10, 145–158.